As the foregoing description demonstrates, the present invention provides effective scratch and date removal techniques for scanned film which may be conveniently implemented using a scanner and software running on a personal computer or other processing device. The techniques of the present invention may also be implemented with hardware components, such as one or more application specific integrated circuits (ASICs), digital signal processors, program-controlled processors, or the like. A combination of software and hardware may also be used to implement the photo extraction technique. With these implementation alternatives in mind, it is to be understood that the block and flow diagrams show the performance of certain specified functions and relationships thereof. The boundaries of these functional blocks have been defined herein for convenience of description. Alternate boundaries may be defined so long as the specified functions are performed and relationships therebetween are appropriately maintained. The diagrams and accompanying description provide the functional information one skilled in the art would require to write program code (i.e., software) or to fabricate circuits (i.e., hardware) to perform the processing required.

While the invention has been described in conjunction with several specific embodiments, many further alternatives, modifications, variations and applications will be apparent to those skilled in the art that in light of the foregoing description. Thus, the invention described herein is intended to embrace all such alternatives, modifications, variations and applications as may fall within the spirit and scope of the appended claims.

What is claimed is:

**1**. A method for processing data in at least one portion of film that has been scanned to generate a pixel representation, comprising the steps of:

(a) segmenting at least one portion of the pixel representation to identify at least one region of missing data;

(b) calculating an area/perimeter ratio for each identified region of missing data; and

(c) subjecting each identified region of missing data having an area/perimeter ratio less than a predetermined maximum to a closest-to-radial-based-function filtering operation to estimate pixel values in that region from neighboring pixel values.

**2**. The method of claim **1**, wherein each pixel represents a color within a three dimensional color space, and wherein the segmenting comprises mapping the three dimensional color space of the pixels in each portion to a one dimensional line segment.

**3**. The method of claim **2**, wherein the segmenting further comprises:

establishing a reference color for each portion; and

determining a corresponding anchor color based on the established reference color;

wherein the one dimensional line segment is defined by the reference color at one end and the anchor color at the other end.

**4**. The method of claim **3**, wherein the segmenting further comprises:

quantizing the one dimensional line segment into a plurality of bins, each of which is identified with a bin-index;

creating a co-occurrence matrix M[i][j] for each portion, M[i][j] being equal to the number of pixel locations in that portion, such that a current pixel has bin-index i and its right or bottom neighbor has bin-index j;

selecting a threshold that creates two areas in the co-occurrence matrix and that maximizes the entropy of the data in each of the two areas; and

identifying pixels having a bin-index greater than the threshold as missing data.

**5**. The method of claim **1**, wherein the calculating comprises performing a component filtering operation.

**6**. The method of claim **1**, wherein, for each pixel value estimated, the closest-to-radial-based-function filtering operation comprises using color values of neighboring pixels, without introducing any new colors, to estimate that pixel value and to fill each region of missing data.

**7**. The method of claim **6**, wherein the closest-to-radial-based-function filtering operation considers spatial distribution and color distribution information in estimating pixel values to fill each region of missing data.

**8**. An apparatus for processing data in at least one portion of a pixel representation of film that has been scanned to generate a pixel representation, comprising:

a segmentation module that segments at least a portion of the pixel representation to identify at least one region of missing data;

a calculator that calculates an area/perimeter ratio for each identified region of missing data; and

a closest-to-radial-based-function filter that subjects each identified region of missing data having an area/perimeter ratio less than a predetermined maximum to a closest-to-radial-based-function filtering operation to estimate pixel values in that region from neighboring pixel values.

**9**. The apparatus of claim **8**, wherein each pixel represents a color within a three dimensional color space, and wherein the segmentation module maps the three dimensional color space of the pixels in each portion to a one dimensional line segment.

**10**. The apparatus of claim **9**, wherein the segmentation module

establishes a reference color for each portion; and

determines a corresponding anchor color based on the established reference color;

wherein the one dimensional line segment is defined by the reference color at one end and the anchor color at the other end.

**11**. The apparatus of claim **10**, wherein the segmentation module

quantizes the one dimensional line segment into a plurality of bins, each of which is identified with a bin-index;

creates a co-occurrence matrix M[i][j] for each portion, M[i][j] being equal to the number of pixel locations in that portion, such that a current pixel has bin-index i and its right or bottom neighbor has bin-index j;

selects a threshold that creates two areas in the co-occurrence matrix and that maximizes the entropy of the data in each of the two areas; and

identifies pixels having a bin-index greater than the threshold as missing data.

**12**. The apparatus of claim **8**, wherein the calculator performs a component filtering operation.

**13**. The apparatus of claim **8**, wherein, for each pixel value estimated, the closest-to-radial-based-function filtering operation comprises using color values of neighboring pixels, without introducing any new colors, to estimate that pixel value and to fill each region of missing data.

**14**. The apparatus of claim **13**, wherein the closest-to-radial-based-function filtering considers spatial distribution and color distribution information in estimating pixel values to fill each region of missing data.

**15**. A machine-readable medium having a program of instructions for directing a machine to process data in at least